UNCLASSIFIED

(2)

AD-A221 420

DTIC ELECTE MAY 0 3 1990 B D

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 90-0385
6a. NAME OF PERFORMING ORGANIZATION The Pennsylvania State University	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research
6c. ADDRESS (City, State, and ZIP Code) Department of Computer Science University Park, PA 16802	7b. ADDRESS (City, State, and ZIP Code) AFOSR/NE, Bldg 410, Bolling AFB, DC 20332-6448.
8a. NAME OF FUNDING SPONSORING ORGANIZATION Air Force Office of Scientific Research	9. PROCUREMENT INSTRUMENT ID NUMBER AFOSR 89-0168
8c. ADDRESS (City, State, and ZIP Code) AFOSR/NE, Bldg 410, Bolling AFB, DC 20332-6448.	10. SOURCE OF FUNDING NUMBERS 3842/A2

11. TITLE *(Include Security Classification)*
Practical Issues in the Complexity of Neural Networks (Unclassified)

12. PERSONAL AUTHOR(S)
I. Parberry, P. Berman, and G. Schnitger

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT	15. PAGE COUNT
Final Technical	88/12/1 to 89/11/30	1990, January 29	7

16. SUPPLEMENTARY NOTATION
The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either express or implied, of the Air Force Office of Scientific Research or the U.S. Government.

17. COSATI CODES	18. SUBJECT TERMS
	Neural networks, complexity theory, fault tolerance, learning.

19. ABSTRACT
The equipment purchased under this Grant was used to supplement the theoretical work done under AFOSR-87-0400 with experimental results. The primary use of the equipment was to perform experiments to aid in the generation and testing of theoretical hypotheses about neural networks, regarding the magnitude of synaptic weights, convergence of learning algorithms, computation and learning with bounded-precision analog neural networks, the performance of simulated annealing on structured problems, and the management of replicated data bases. Research is still underway to gather more experimental data and provide theoretical justification for the observations.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT Unclassified	21. ABSTRACT SECURITY CLASSIFICATION Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE	22c. OFFICE SYMBOL
Dr. Alan Craig	202-767-4931	NE

UNCLASSIFIED

AFOSR-TR- 90-0385

PRACTICAL ISSUES IN THE COMPLEXITY OF NEURAL NETWORKS

Ian Parberry
Piotr Berman
Georg Schnitger
Dept. of Computer Science
333 Whitmore Laboratory
Penn State University
University Park, Pa 16802.

29 January, 1990

Final Technical Report

Prepared for

Dr. A. Craig
AFOSR/NE, Building 410
Bolling AFB DC 20332-6448

Practical Issues in the Complexity of Neural Networks: Final Technical Report

Ian Parberry
Piotr Berman
Georg Schnitger

Dept. of Computer Science
333 Whitmore Laboratory
Penn State University
University Park, Pa 16802.

1. Research Objectives

One criticism of theoretical neural network research is that the results are of asymptotic interest only, that is, the ''hidden'' constant multiples in the resource analysis make it not of immediate practical utility. Our objectives were to leaven the theoretical work done under Grant AFOSR-87-0400 with practical experience. Experiments were also done in order to generate open questions, hypotheses, and conjectures which are the subject of on-going research.

2. Accomplishments

2.1. Synaptic Weights

The discrete neural network model uses neurons which compute a function of the form $f:\mathbf{B}\to\mathbf{B}$ (where $\mathbf{B}$ denotes the Boolean set $\{0,1\}$), such that

$$f(x_1,\ldots,x_n)=g(\sum_{i=1}^{n} w_i x_i),$$

where g is the step function $g:\mathbf{R}\to\mathbf{B}$ defined by $g(x)=1$ iff $x\geq 0$, and $w_1,\ldots,w_n\in\mathbf{R}$.

Muroga et al [1] showed that the weights can be made integers bounded above in magnitude by $O(n^{n/2})$. Their proof reduces to finding the maximum value of the ratio of the determinants of two zero-one matrices, where the numerator is equal to the denominator with one column replaced by a zero-one vector. They ignore the denominator and bound the numerator. We are currently trying to improve this upper-bound, or to find a matching lower-bound. The only progress that we have made so far is to extend the result to non-Boolean domains.

We have run experiments that lead us to believe that a better worst-case upper-bound is $O(n^{n/4})$, although $2^{O(n)}$ is still possible. Preliminary theoretical results indicate that the average synaptic weight is very small. Our experiments have verified that this is true even for small values of n, indicating that the theory in this case is immediately applicable in practice, not just for asymptotically large values of n that would be impossible to reach with current technology.

Research on this subject is still in progress. We are currently writing a more efficient program which will enable us to evaluate the ratio of the determinants for larger values of n. The initial program was written in Pascal, and consisted of a multi-word arithmetic package, a rational arithmetic package, a matrix manipulation package, and a main body which picked a large number of random matrices and vectors, and evaluated the ratio of the determinants described above. The program was hand-coded, without the use of any proprietary software or libraries. We plan to translate the program to *gcc* to improve its speed. Results will be reported in [1].

A better upper-bound for the synaptic weights would lead to improved upper-bounds for many problems, such as integer addition, and multiplication. These functions can be computed in constant depth and polynomial size by a discrete neural network; the size would be reduced by a polylog factor. It would also reduce the size required for the simulation of a discrete neural network by one with unit weights.

2.2. Convergence of Learning Algorithms

We investigated the speed of convergence for distributed learning algorithms learning a threshold function on a single threshold gate. Among the considered learning algorithms are Perceptron learning, Littlestone's Winnow algorithm, the delta rule and the generalized delta rule. The results so far are inconclusive, but indicate slow convergence of all mentioned procedures in general. The emphasis of further research will be on finding specific examples on which Littlestone's procedure performs poorly (his procedure will learn thresholds with small weights and relatively large separation).

2.3. Analog Neural Networks

The equipment was also used for the design and testing of some of the neural circuits reported in [2,3]. These papers analyze the complexity of *computation and learning* in analog neural networks with limited precision.

2.4. Simulated Annealing

There has been much interest recently in using neural networks to solve optimization problems stochastically using Simulated Annealing (for example, the Boltzmann machine). It is not clear that Simulated Annealing does well on examples generated in practice. In a class on Computer Aided Design for VLSI, students did research on Simulated Annealing and on general purpose routing procedures (line expansion, computational geometry approaches). The emphasis in the simulated annealing project was a comparison of Simulated Annealing with special purpose procedures (Kernighan-Lin; Leighton-Rao) for Graph Bisection. Whereas it is known that Simulated Annealing performs well (given sufficient time) for random graphs, little work has apparently been done for highly structured graphs. The results of the project concern the hypercube, cube-connected cycles and weakly-connected graphs (a singly interconnected collection of rings). Simulated Annealing did not perform significantly better than a randomized version of Kernighan-Lin, but used significantly more time. We hope to extend this study later to cover a broader class of highly structured graphs.

2.5. Distributed Data Bases

In a replicated data base one can increase availability of data by allowing the retrieval or modification of the data even if not all copies are accessible. This way transactions may be performed even when some nodes (computers) or links are not operational. The problem is to maximize the availability while preventing the existence if inconsistent copies of the data (preserving the integrity of the replicated data base).

One scheme is to assign each node a weight (number of votes according) to its reliability and importance for the network, and perform transactions only if the nodes possessing the absolute majority of the votes are available. In general, there exist networks where the voting scheme cannot yield optimal availability. The open question, which was subject of extensive testing, is the following: can one use the voting scheme to approximate well the optimal schema of replica control? The results suggest that the the voting scheme is not significantly inferior from the optimal solution (while always much simpler and easier to implement).

The second subject of testing were heuristics to learn an optimal or near-optimal assignment of votes, using the methodology of neural networks. Certain families of special cases were used as benchmark. Regretfully, no heuristic handled all benchmark cases in a satisfactory manner. The difficulty is that, unlike in a classic neural network, the heuristic do not have precise feedback. The empirical vote assignment should result in allowing transactions with maximal possible frequency, subject to integrity constraints. However, the mere fact that a transaction is not allowed is not a sufficient basis for a negative feedback. We are still trying to find a proper correction for negative feedbacks and a proper regime of weight updates.

The more theoretical part of this research resulted in a technical report "Voting and Other Static Schemes for Managing Replicated Data Bases" P. Berman, M. Obradovic, Technical Report CS-89-46, Department of Computer Science, Penn State University, 1989.

2.6. Related Research

In a class on Computer Graphics and Computational Geometry, the equipment was used mainly to experiment with fractals. This included work on rewriting systems, Julia sets and the Mandelbrot as well as random fractals for modeling mountains. We would have been unable to perform these projects with departmental equipment due to the computation intensive algorithms, particularly for Julia sets and the Mandelbrot Set. Besides enabling us to give proper instruction, these projects also brought up research questions that we intend to work on in the future. These questions are mainly concerned with the computational complexity of algorithms for fractals, an area that apparently received little attention in the past.

3. Conclusion

The equipment grant was used to purchase a moderate-cost state-of-the-art computing environment for the Principal Investigators and Research Assistants of Grant AFOSR-

87-0400. In addition to the experiments performed in conjunction with that research, the equipment provided the daily computing environment for the researchers, performing such necessary tasks as electronic mail communication with researchers at other Institutions, and the typesetting of research publications. The Department of Computer Science at Penn State University was and is financially unable to provide adequate computing facilities to faculty. The equipment has done much to increase the productivity and quality of the research of all concerned. These secondary benefits are not to be ignored.

4. Publications

1. I. Parberry, The Computational and Learning Complexity of Neural Networks, In Preparation.

2. Z. Obradovic and I. Parberry, ''Analog neural networks of limited precision I: Computing with multilinear threshold functions (preliminary version)'', Proc. 1989 IEEE Conference on Neural Information Processing Systems, Denver, Co, To Appear.

3 Z. Obradovic and I. Parberry, ''Analog neural networks of limited precision II: Learning with multilinear threshold functions preliminary version)'', to be submitted to COLT '90.

5. Personnel

Name	*Position*
Dr. Ian Parberry Dr. Georg Schnitger Dr. Piotr Berman	Principal Investigator Co-Principal Investigator Co-Principal Investigator
Mirjana Obradovic Zoran Obradovic Sanjeev Dharap	Research Assistant for AFOSR-87-0400 Research Assistant for AFOSR-87-0400 Student

6. Conference Presentations

Z. Obradovic and I. Parberry, "Analog neural networks of limited precision I: Computing with multilinear threshold functions (preliminary version)", Presented by the co-authors at a Poster Session of the 1989 IEEE Conference on Neural Information Processing Systems, Denver, Co, November 1989.

7. References

1. S. Muroga, I. Toda, and S. Takasu, "Theory of majority decision elements," *J. Franklin Inst.*, vol. 271, pp. 376-418, May 1961.